UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

DEFENSE TECHNICAL INFORMATION CENTER

a264576

REPORT DOC

DTIC
S ELECTE
MAY 2 0 1993
C D

1a. REPORT SECURITY CLASSIFICATION Unclassified	
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) #65
6a. NAME OF PERFORMING ORGANIZATION Institute for Brain and Neural Systems 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Personnel and Training Research Programs Office of Naval Research (Code 1142PT)
6c. ADDRESS (City, State, and ZIP Code) Brown University Providence, Rhode Island 02912	7b. ADDRESS (City, State, and ZIP Code) 800 North Quincy Street Arlington, VA 22217-5000
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-91-J-1316
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO / PROJECT NO / TASK NO / WORK UNIT ACCESSION NO

11. TITLE (Include Security Classification)

Three Dimensional Object Recognition Using an Unsupervised BCM Network: The Usefulness of Distinguishing Features.

12. PERSONAL AUTHOR(S)

Nathan Intrator and Joshua I. Gold

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical Report	FROM ______ TO ______	May 17, 1993	Twelve

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	BCM, Feature Extrac ion, Unsupervised Object Recognition
05	08		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

We propose an object recognition scheme based on a method for feature extraction from gray level images that corresponds to recent statistical theory, called projection pursuit, and is derived from a biologically motivated feature extracting neuron. To evaluate the performance of this method we use a set of very detailed psychophysical 3D object recognition experiments (Bulthoff and Edelman, 1992).

93 5

93-11613

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Joel Davis	22b. TELEPHONE (Include Area Code) (703) 696-4744 22c. OFFICE SYMBOL

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

Three-Dimensional Object Recognition Using an Unsupervised BCM Network: The Usefulness of Distinguishing Features*

Nathan Intrator† **Joshua I. Gold**
Center for Neural Science
Box 1843, Brown University
Providence, RI 02912

May 26, 1992

Abstract

We propose an object recognition scheme based on a method for feature extraction from gray level images that corresponds to recent statistical theory, called projection pursuit, and is derived from a biologically motivated feature extracting neuron. To evaluate the performance of this method we use a set of very detailed psychophysical 3D object recognition experiments (Bülthoff and Edelman, 1992).

Keywords: unsupervised learning, feature extraction, dimensionality reduction, object recognition.

*To Appear in *Neural Computation*
†Research was supported by the National Science Foundation, the Army Research Office, and the Office of Naval Research.

1 Introduction

A system which performs recognition of three-dimensional objects in visual space must transform a complex pattern of visual inputs to an appropriate categorization. Such recognition is possible, for example, by template matching once the object and its templates are brought into register (Ullman, 1989). Other similar schemes (Lowe, 1986; Thompson and Mundy, 1987) base the recognition on viewpoint consistency, which relate projected locations of key features of a model to its 3D structure given a hypothesized view point. The regularization network or HyperBF interpolation scheme (Poggio and Edelman, 1990; Poggio and Girosi, 1990) represents 3D objects by sets of 2D views using vectors of key-feature locations and regards generalization from familiar to novel views as a problem of nonlinear hypersurface interpolation in the space of all possible views. All these methods rely on the ability to find key features in the objects and in some cases, to solve the correspondence problem between them.[1] However, sometimes these tasks can be as difficult as the recognition itself.

In this paper, we propose an object recognition method that does not rely on finding such key features a-priori. Instead, a transformation is sought which reduces the pixel image representations into a low dimensional space from which nonlinear hypersurface interpolation can be used for the recognition task. The dimensionality reducing transformation is based on projecting the pixel image onto a set of *object features*. The actual form of object features, and methods of extracting them, are not at all clear and are subject to current research in many disciplines (Edelman, 1991). We propose to use a method for feature extraction which corresponds to recent statistical theory (Friedman and Tukey, 1974; Friedman, 1987) and is based on a biologically motivated feature extracting neuron. To evaluate the performance of this method based on the above criteria we use a set of very detailed psychophysical 3D object recognition experiments (Bülthoff and Edelman, 1992). These psychophysical experiments were specifically constructed to test several theories of object representation and recognition and are therefore appropriate for testing the usefulness of our features of recognition.

2 A new model for object recognition based on a novel set of features

Many feature extraction theories for object recognition are based on the assumption that objects are represented by clusters of points in a high dimensional feature space. (Duda and Hart, 1973). However, finding clusters in very high dimensional space suffers from the inherent sparsity of such space, and therefore can not be directly approached by classical methods such as cluster analysis (Duda and Hart, 1973), discriminant analysis (Fisher, 1936; Sebestyen, 1962), or factor analysis (Harman, 1967).

Recent work (Intrator, 1990; Intrator and Cooper, 1992) connecting biologically motivated feature extraction networks (Bienenstock et al., 1982, henceforth to be referred to as "BCM") with sophisticated statistical techniques (Friedman and Tukey, 1974; Friedman, 1987) suggests that this

[1] Edelman and Weinshall (1991) used the vertices without solving the correspondence problem between them.

problem may be approached by extending a method known as Exploratory Projection Pursuit. Use of this method provides both a rigorous mathematical definition of "salient features of recognition" and a procedure for extracting them.

2.1 Feature Extraction in High Dimensional Space – the BCM Model

From a mathematical viewpoint, extracting features from gray level images is related to dimensionality reduction in high dimensional vector space, in which an $n \times k$ pixel image is considered to be a vector of length $n \times k$. The *curse of dimensionality* (Bellman, 1961) says that it is impossible to base the recognition on the high dimensional vectors directly, because the number of training patterns needed for training such a classifier should increase in an exponential order with the dimensionality. Thus, if the *important structure* (for classification) can be represented in a low dimensional space, dimensionality reduction should take place before attempting the classification. Furthermore, due to the large number of parameters involved, a feature extraction method that uses the class labels of the data may miss important structure that is not exhibited in the class labels, and therefore be more biased to the training data than a feature extractor that relies on the high dimensional structure (Huber, 1985). This suggests that an unsupervised feature extraction method may have better generalization properties in high dimensional problems.

In this paper we concentrate on a specific form of unsupervised dimensionality reduction/feature extraction. This form relies on the notion of distinguishing features which focus on discrimination among classes and not faithful representation of the data. Thus, this form is different[2] from classical methods such as factor analysis (Harman, 1967, for review) which tend to combine features that seem to have high correlation, or principal component analysis which seeks directions that maximize the variance of the projected distribution.[3]

A general framework for feature extraction is Projection Pursuit, and its unsupervised version - Exploratory Projection Pursuit (Kruskal, 1969; Friedman and Tukey, 1974; Friedman, 1987; Huber, 1985, for review). The idea behind projection pursuit is to pick *interesting* low dimensional projections of a high dimensional point cloud by maximizing an objective function called the projection index. The projection index usually measures some form of deviation from normality of the projected distribution.[4] Intrator (1990) presented a multiple feature extraction method that seeks multi-modality in the projected distributions. This method is based on a modified version of the BCM neuron (Bienenstock et al., 1982). The biological relevance of this neuron has been extensively studied (Bear et al., 1987; Bear and Cooper, 1990; Gold, 1991), and it was shown that results of this method are in agreement with classical visual deprivation experiments (Clothiaux et al., 1991). Sets of these neurons which are organized in a lateral inhibition architecture (Intrator, 1990; Intrator and Cooper, 1992), which forces different neurons in the network to find different projections (i.e., features), combined with the simplicity of the projection index, make this method computationally practical for multiple feature extraction in high dimensional spaces

[2]See (Intrator and Cooper, 1992) for a discussion of the difference

[3]Principal components may not retain enough structure needed for classification (Duda and Hart, 1973, p. 212)

[4]For a discussion on various projection indices, see Huber (1985), Jones and Sibson (1987), Intrator and Cooper (1992).

(Intrator, 1992).

3 Application of the Model to 3D Object Recognition

The combined unsupervised feature extraction/classification method used in these experiments is described in Intrator (1992). In general, the generalization properties of hybrid feature extraction/classification method depend on the feature extraction as well as the classification method used. Edelman and Poggio (1990) have attempted the recognition of the same 3D wire-like objects discussed in this paper, by extracting a priori an ordered list of vertices from the image and using a generalized radial basis function classification scheme (Moody and Darken, 1989; Poggio and Girosi, 1990, GRBF). This method classified lists of vertices based on their orientation within a vector space defined by the vertex sets of known objects; it achieved close to human performance in generalizing to novel views of the wires. The performance reflected a strong focus on the classification technique, and assumed a deterministic, a-priori feature extraction. We, on the other hand, want to concentrate on the examination of the properties of our proposed feature extraction method and therefore in this study have chosen to use a classical, well-known classifier, based on the *k-nearest-neighbor-rule*[5] (see for example, Duda and Hart, 1973).

In addition to the type of classifier used, the recognition paradigm with which the system is tested is a vital component in determining the usefulness of the features extracted. In the following sections we present an application of the BCM model to a set of specific 3D object recognition problems. The experiments chosen fulfill two important criteria: 1) they test the model's abilities to both recognize and generalize across a wide range of difficulties, and 2) these same studies have been used to test the abilities of not only computational models, but also human subjects; the psychophysical results in fact serve as benchmarks for this study.

3.1 Previous Studies

Edelman and Bülthoff (1990, 1991) developed and used wire-like objects in their experiments, in an effort to simplify the problem for the feature-extractor by providing little or no occlusion of the key features from any viewpoint. The wires consisted of seven connected segments, each pointed in a random direction but with its vertices distributed normally around the origin. Each experiment consisted of two phases, training and testing. In the training phase subjects were shown the target object from two standard views, located 75 degrees apart along the equator of the viewing sphere. The target oscillated around each of the two standard orientations with an amplitude of ± 15 degrees about a fixed vertical axis, with views spaced at 3-degree increments (see Figure 1). Test views were located either along the equator – on the minor arc bounded by the two standard views (INTER condition) or on the corresponding major arc (EXTRA condition) – or on the meridian passing through one of the standard views (ORTHO condition). Testing was conducted according to a two-alternative forced choice (2AFC) paradigm, in which subjects were asked to indicate whether the

[5]Very similar classification results where obtained using a back-propagation classifier. In a forthcoming article, performance of back-propagation and radial basis function (RBF) classifiers will be compared using features extracted by the above feature extraction method.

displayed image constituted a view of the target object shown during the preceding training session. Test images were either unfamiliar views of the training object or random views of a distractor (one of a distinct set of objects generated by the same procedure).

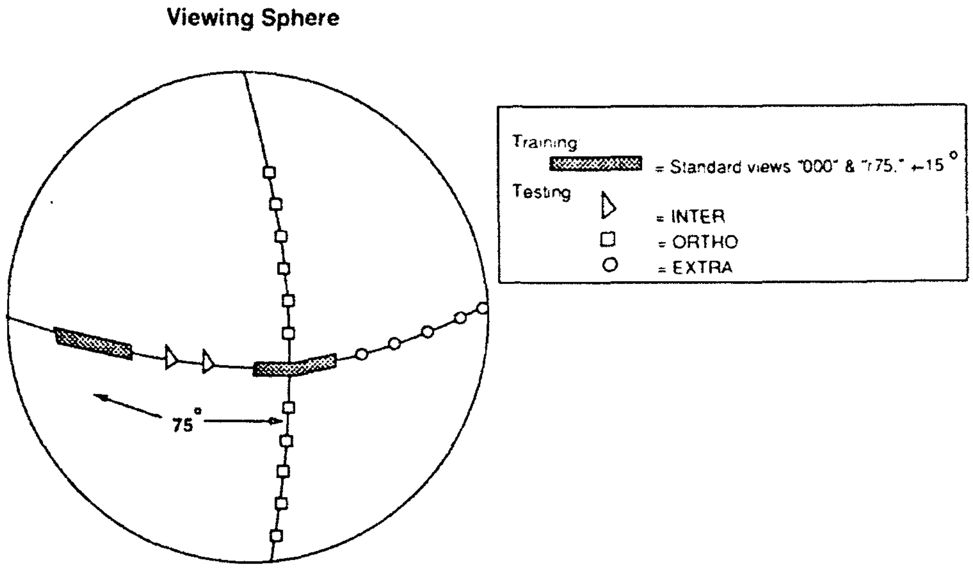

Figure 1: The training and testing experimental paradigm

A number of interesting characteristics of human visual object recognition abilities emerged from the psychophysical experiments. Generalization over orientations lying between two sets of known views - the INTER condition - resulted in, on average, significantly fewer errors than with the other two extrapolation conditions. In addition, error rates increased steadily as the testing views moved farther away from the learned views, until recognition was near chance levels at large displacements. Finally, there were indications for a "horizontal bias," so that error rates were lower when generalization was required along the horizontal, as opposed to the vertical, plane.

3.2 Experimental Paradigm

In the first part of the study, the network was tested on a 63 by 63 array of 8-bit gray-scale values with a paradigm nearly identical to the one used in the psychophysical investigation (Edelman and Bülthoff, 1991). The procedure was modified slightly in that training was performed with two wires, since the k-NN classifier would yield meaningless results if trained on only a single wire.

In the second part of the study, simple yes/no recognition was upgraded to a more difficult classification task involving six separate wires. The modification was necessary in order to fully test the BCM model's ability to extract the most salient rotation-invariant features from the images. Specifically, since BCM neurons explicitly search for differentiating features (due to the search for multi-modality in the projected distribution), many cases involving only two distinct sets of inputs can be solved with "features" corresponding to prototypical views of each wire. In these cases, the two sets of wire-views, corresponding to the two wires, would form two distinct clusters in feature space. However, such differentiation would be much more difficult with a larger number of wires, and therefore the BCM network neurons would be forced to find projections that correspond to individual, rotation-invariant features, not prototypical views of individual wires.

In addition, the model was modified in an attempt to account for the asymmetric psychophysical results. In the psychophysical experiments, the horizontal bias was found when humans were given the *exact* same paradigm as described above, except the objects were rotated 90 degrees so that the training axis was aligned vertically, not horizontally. One possible explanation of such asymmetry is in increased resolution at the object representation level, namely, due to the fact that behaviorally, humans spend more time rotating around a vertical axis (i.e., rotation in a horizontal plane). This is experimentally equivalent to having more patterns rotated in a horizontal than in a vertical plane. This possibility has been eliminated in the careful psychophysical experiment performed by Edelman and Bülthoff (1991), in which subjects are provided identical experience with horizontal and vertical training. The continued existence of the bias under such conditions implicates an internal mechanism. We hypothesized greater a-priori resolution in the *internal* representation along the horizontal plane.[6] More specifically, we set the ratio between the resolution in the horizontal plane and that in the vertical plane (the aspect ratio) to be 2/1 for "normal" training in the horizontal plane; conversely, training in the vertical plane was, from the point of view of the network, equivalent to setting the aspect ratio to be 1/2. Prediction of simulation performance due to this asymmetrical resolution is not straightforward since there are two contradictory effects. On the one hand, decreased resolution in the vertical plane means reduced disparity from rotations along that plane and therefore possibly better performance. On the other hand, there may also be improved performance in the horizontal axis since higher resolution will emphasize features which are rotation invariant along that direction.

4 Results

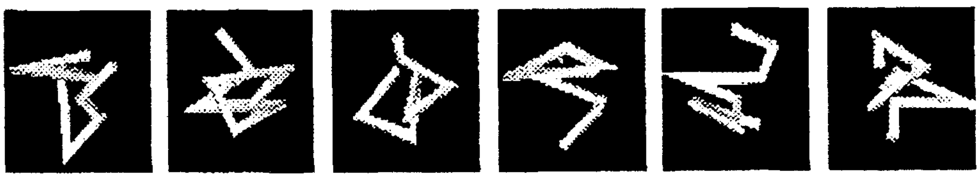

Figure 2: The six wires from a single view.

The 6 wires used in the experiments are depicted in Figure 2. Different views of three of the wires are depicted in Figure 3. When only two wires were used (experiment one) the features extracted correspond almost exclusively to a single view of a whole image of one of the wires.

In contrast, when the task was recognition of six wires the extracted features emphasized small patches of several images or views, namely, areas that either remain relatively invariant under the

[6]There is, in fact, limited evidence for visual field elongation in the horizontal plane (Hughes, 1977).

rotation performed during training or represented a major differentiating characteristic of a specific wire (Figure 4). A typical result is a set of weights that may correspond to a single wire but emphasizes small patches of the object and selectively inhibits selected areas which correspond to invariant locations of adjacent wires. For example, the top left image of Figure 4 largely represents object number 5 in Figure 2 with additional inhibition from other objects, while the top right image or the bottom second from the right image exhibit weights related to several images/views.

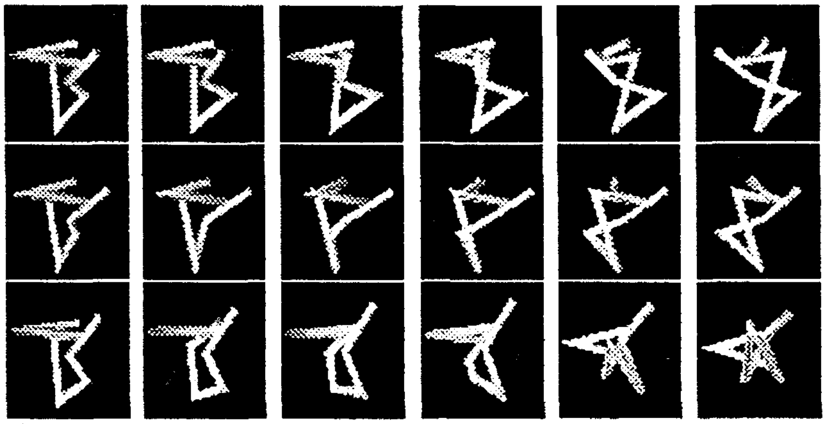

Figure 3: Different views (15 degrees apart) of a single wire; top-to-bottom are INTER, EXTRA, and ORTHO.

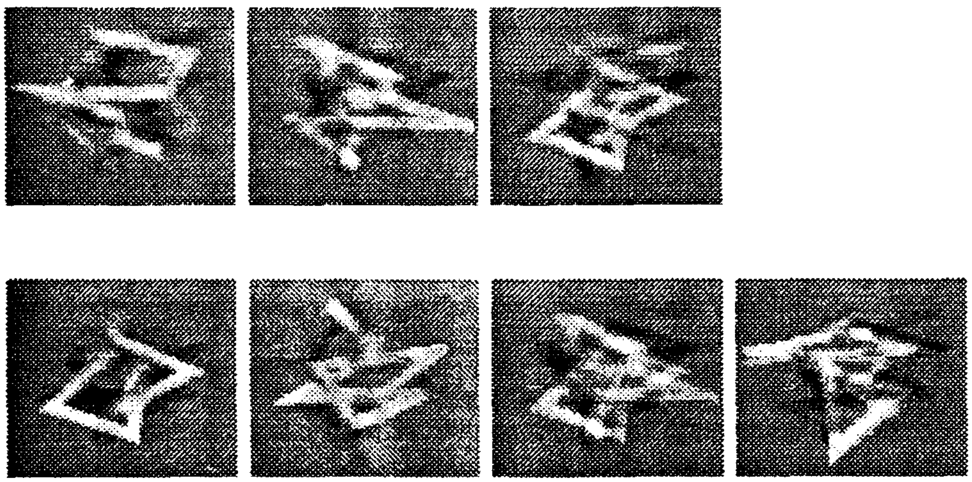

Figure 4: Rotation invariant features for tube-like objects extracted using a network of 7 BCM neurons trained on 6 tube-like objects. White areas represent strong synaptic weights, black areas represent negative synaptic weights (inhibition).

Classification results demonstrate the usefulness of the extracted features: generalization in the INTER orientations resulted in consistently low error rates – around 15% (in which the chance error rate on this six wire experiment is 83.3%) – which indicates that the features extracted by the BCM network could generalize well in those new views.[7] Furthermore, the results are comparable to those obtained in the psychophysical experiments. First, INTER recognition resulted in, on average, significantly fewer errors than with the other two extrapolation conditions. Second, error rates increased steadily as the testing views moved farther away from the learned views, until recognition was near chance levels at large displacements. These results are analogous to the ones shown in Figure 5 in which the aspect ratio is 2/1.

Taken together, Figures 5 and 6 demonstrate a horizontal bias as seen in the psychophysical studies. When aspect ratio is 0.5, which corresponds in our model to training on rotations in the vertical plane, INTER performance is worse. This result suggests that finding specific rotation invariant features was harder in that case, given its lower resolution. On the other hand, there is no significant change in the performance of EXTRA and ORTHO orientations, suggesting that the extracted features were in both situations equally useful for EXTRA and ORTHO orientations.

Figures 7 and 8 show the results of the experimental paradigm testing the effect of additional experience during training in the horizontal plane.[8] Both figures show results on training with an aspect ratio of 1, i.e., no resolution asymmetry was used between the horizontal and vertical plane. In the experiments summarized in Figure 7, the same number of training views (experience) as in the previous set of experiments were used. In the experiments summarized in Figure 8, half as many training views were used. A number of interesting observations can be made. Results on the INTER condition for an aspect ratio of 1 behave as can be predicted from the previous set of experiments; specifically, error rates were in between those of aspect ratios 2 and 0.5. EXTRA and ORTHO results, however, were less noticeably affected, indicating that object resolution primarily affected the discovery of rotation invariant features to be used for recognition in the INTER condition, as opposed to reducing overall recognition ability. Results from Figure 8,however, demonstrate a different effect. Reducing the number of training patterns, analogous to reducing the experience of vertical training, does not lead to an asymmetry in specific recognition conditions, but instead to a general decline in overall recognition ability. This suggests that reducing the number of training views in a model (. .thout reducing the overall training angle rotation) does not simply affect the ability to extract rotation-invariant features for a particular recognition task. Instead, it degrades the ability of the model in overall feature extraction performance.

5 Discussion

This paper touches on issues of object representation. It is assumed that an object is internally represented by a particular combination of features. The nature of these features and the means for binding together the most important combination of features are still undetermined (Sejnowski, 1986). We presented an unsupervised method for extracting features directly from grey level pixel images, and we showed that a surprisingly small number of features is needed for a complex clas-

[7]Additional support to the usefulness of the extracted features to rotation invariant recognition is shown in a subsequent work (Intrator et al., 1991; Sklar et al., 1991) in which the extracted features are used to occluded parts of the images and another network is trained to recognize the occluded images.

[8]Testing in both cases used the same number of patterns as in the previous experiments.

sification task. A comparison of our results to similar psychophysical experiments gives some indication that these features posses desired invariance properties which allow for overall classification performance that compares favorably with human performance.

Extracting features from these gray level images is a highly non-trivial statistical task. The dimensionality of this problem is 63×63 pixels; therefore, the *curse of dimensionality* implies that the number of training patterns should be immense, and yet from a small training set of 132 wires useful directions (projections) were extracted corresponding to features which were especially useful for rotation invariant recognition. This suggests that the BCM network may be a practical tool for gray level image recognition in which internal low dimensional feature representation emerges as a result of unsupervised training.

Acknowledgements

We wish to thank Heinrich Bülthoff and Shimon Edelman for the encouragement and many fruitful conversations that have led to this paper. Dave Sheinberg, Phillipe Schyns and Eric Sklar were invaluable for their help in using the AVS^{TM} system for getting the gray-level images. Finally, the excellent computational facilities of the Cognitive Science Department at Brown University allowed us to complete the simulations required for this project.

Research was supported by the National Science Foundation, the Army Research Office, and the Office of Naval Research.

References

Bear, M. F. and Cooper, L. N. (1990). Molecular mechanisms for synaptic modification in the visual cortex: Interaction between theory and experiment. In Gluck, M. and Rumelhart, D., editors, *Neuroscience and Connectionist Theory*, pages 65–94. Lawrence Erlbaum, Hillsdale, New Jersey.

Bear, M. F., Cooper, L. N., and Ebner, F. F. (1987). A physiological basis for a theory of synapse modification. *Science*, 237:42–48.

Bienenstock, E. L., Cooper, L. N., and Munro, P. W. (1982). Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *Journal Neuroscience*, 2:32–48.

Bülthoff, H. H. and Edelman, S. (1992). Psychophysical support for a 2-D view interpolation theory of object recognition. *Proceedings of the National Academy of Science*, 89:60–64.

Clothiaux, E. E., Cooper, L. N., and Bear, M. F. (1991). Synaptic plasticity in visual cortex: Comparison of theory with experiment. *Journal of Neurophysiology*, 66:1785–1804.

Duda, R. O. and Hart, P. E. (1973). *Pattern Classification and Scene Analysis.* John Wiley, New York.

Edelman, S. (1991). Features of recognition. CS-TR 10, Weizmann Institute of Science.

Edelman, S. and Bülthoff, H. H. (1991). Orientation dependence in the recognition of familiar and novel views of 3D objects. *Vision Research.* submitted.

Edelman, S. and Poggio, T. (1990). Bringing the Grandmother back into the picture: a memory-based view of object recognition. A.I. Memo No. 1181, Artificial Intelligence Laboratory, Massachusetts Institute of Technology. to appear in Int. J. Pattern Recog. Artif. Intell.

Edelman, S. and Weinshall, D. (1991). A self-organizing multiple-view representation of 3D objects. *Biological Cybernetics*, 64:209–219.

Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7:179–188.

Friedman, J. H. (1987). Exploratory projection pursuit. *Journal of the American Statistical Association*, 82:249–266.

Friedman, J. H. and Tukey, J. W. (1974). A projection pursuit algorithm for exploratory data analysis. *IEEE Transactions on Computers*, C(23):881–889.

Gold, J. I. (1991). A model of dendritic spine head [Ca^{++}]: Exploring the biological mechanisms underlying a theory for synaptic plasticity. Unpublished honors thesis, Brown University.

Harman, H. H. (1967). *Modern Factor Analysis*. University of Chicago Press, Second Edition, Chicago and London.

Huber, P. J. (1985). Projection pursuit. (with discussion). *The Annals of Statistics*, 13:435–475.

Hughes, A. (1977). The topography of vision in mammals of contrasting live style: Comparative optics and retinal organisation. In Crescitelli, F., editor, *The Visual System in Vertebrates, Handbook of Sensory Physiology VII/5*, pages 613–756. Springer Verlag, Berlin.

Intrator, N. (1990). A neural network for feature extraction. In Touretzky, D. S. and Lippmann, R. P., editors, *Advances in Neural Information Processing Systems*, volume 2, pages 719–726. Morgan Kaufmann, San Mateo, CA.

Intrator, N. (1992). Feature extraction using an unsupervised neural network. *Neural Computation*, 4:98–107.

Intrator, N. and Cooper, L. N. (1992). Objective function formulation of the BCM theory of visual cortical plasticity: Statistical connections, stability conditions. *Neural Networks*, 5:3–17.

Intrator, N., Gold, J. I., Bülthoff, H. H., and Edelman, S. (1991). Three-dimensional object recognition using an unsupervised neural network: Understanding the distinguishing features. In Feldman, Y. and Bruckstein, A., editors, *Proceedings of the 8th Israeli Conference on AICV*, pages 113–123. Elsevier.

Jones, M. C. and Sibson, R. (1987). What is projection pursuit? (with discussion). *J. Roy. Statist. Soc.*, Ser. A(150):1–36.

Kruskal, J. B. (1969). Toward a practical method which helps uncover the structure of the set of multivariate observations by finding the linear transformation which optimizes a new 'index of condensation'. In Milton, R. C. and Nelder, J. A., editors, *Statistical Computation*. Academic Press, New York.

Lowe, D. G. (1986). *Perceptual organization and visual recognition*. Kluwer Academic Publishers, Boston, MA.

Moody, J. and Darken, C. (1989). Fast learning in networks of locally tuned processing units. *Neural Computation*, 1:281–289.

Poggio, T. and Edelman, S. (1990). A network that learns to recognize three-dimensional objects. *Nature*, 343:263–266.

Poggio, T. and Girosi, F. (1990). Networks for approximation and learning. *IEEE Proceedings*, 78(9):1481–1497.

Sebestyen, G. (1962). *Decision Making Processes in Pattern Recognition*. Macmillan, New York.

Sejnowski, T. J. (1986). Open questions about computation in Cerebral Cortex. In McClelland, J. L. and Rumelhart, D. E., editors, *Parallel Distributed Processing*, volume 2, pages 372–389. MIT Press, Cambridge, MA.

Sklar, E., Intrator, N., Gold, J. J., Edelman, S. Y., and Bülthoff, H. H. (1991). A hierarchical model for 3D object recognition based on 2D visual representation. In *Neurosci. Soc. Abs.*

Thompson, D. W. and Mundy, J. L. (1987). Three-dimensional model matching from an unconstrained viewpoint. In *Proceedings of IEEE Conference on Robotics and Automation*, pages 208–220, Raleigh, NC.

Ullman, S. (1989). Aligning pictoral descriptions: an approach to object recognition. *Cognition*, 13:13 – 254.

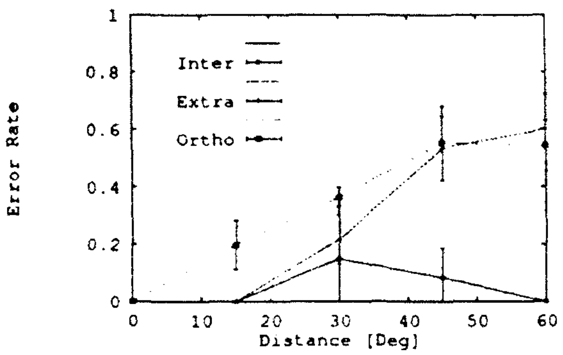

Figure 5: Fraction of misclassification performance for wires trained on the horizontal plane.

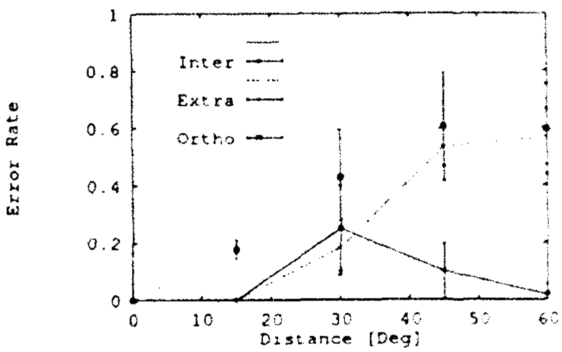

Figure 6: Fraction of misclassification performance for wires trained on the vertical plane. *Note the degradation in performance in the* INTER orientations.

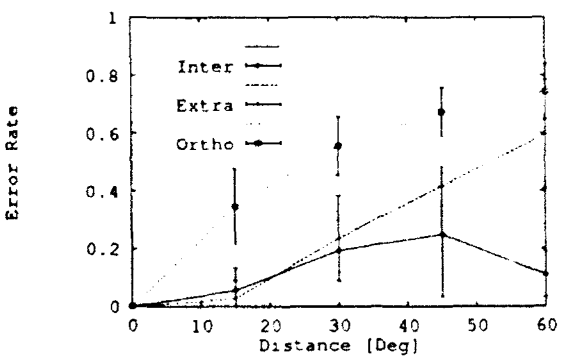

Figure 7: Fraction of misclassification performance for wires trained on the horizontal plane with no asymmetry.

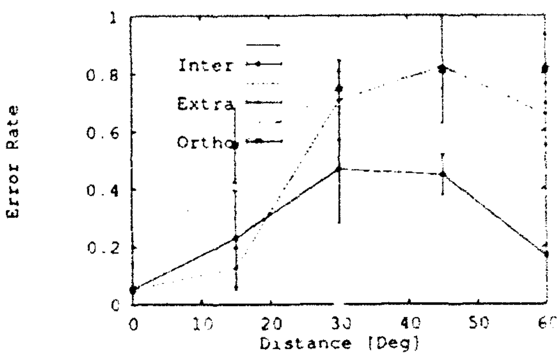

Figure 8: Fraction of misclassification performance for wires trained with reduced training experience (views).